

# Master Data Management & Big Data Analytics

Potential WaTech Services



"the consolidated technology services agency -RCW 43.105.006"

# Master Data Management (MDM)



# Master Data: Gartner definition

“...the consistent and uniform set of identifiers and extended attributes that describes the core entities of the enterprise including

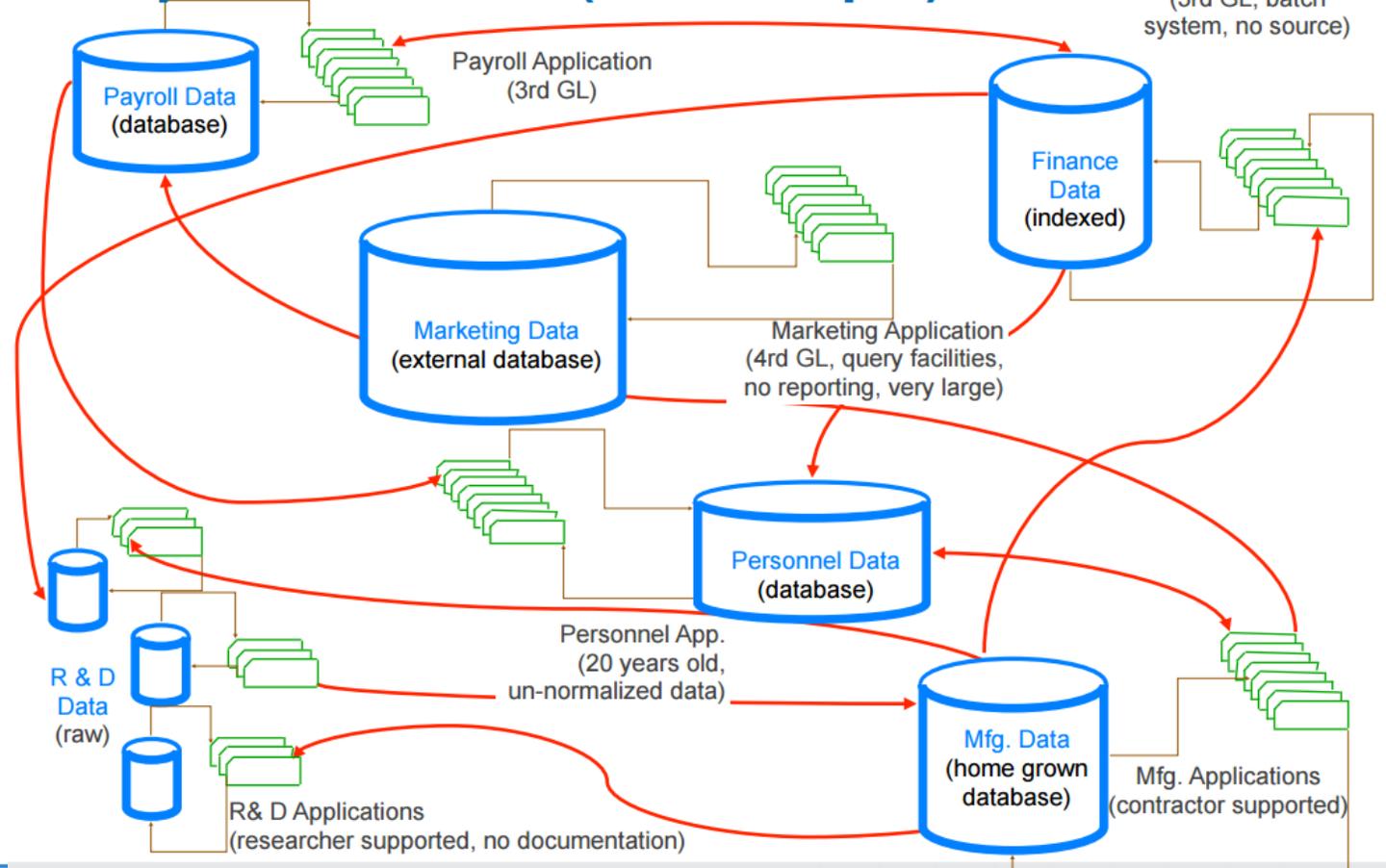
- **customers,**
- **prospects,**
- **citizens,**
- **suppliers,**
- **sites,**
- **hierarchies, and**
- **chart of accounts.”**



# Historical Approaches

- Master Files
- Reference Data
- Look-up Tables
- Point-to-Point Interfaces

## Multiple Sources of (for example) Customer Data

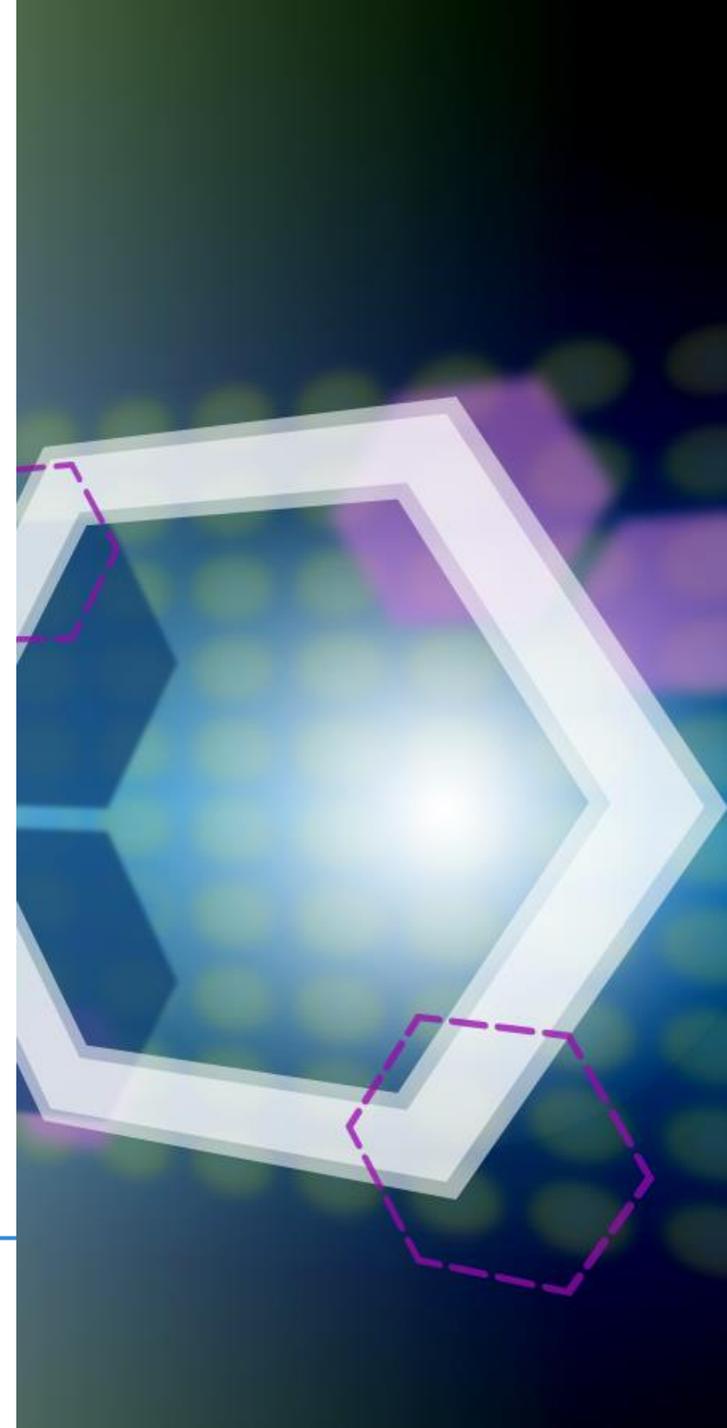


# Master Data Management (MDM): Gartner definition

“... a technology-enabled discipline in which business and IT work together to ensure the

- **uniformity,**
- **accuracy,**
- **stewardship,**
- **semantic consistency and**
- **accountability**

of the enterprise’s official shared master data assets.”



# MDM Technology Approaches

## Do It Yourself (DIY)

- ◆ Purely code and database design
- ◆ Hybrid, e.g. with Microsoft MDM

## Domain Specific MDM products

- ◆ Customer Data Integration (CDI)
- ◆ Product Information Management (PIM)

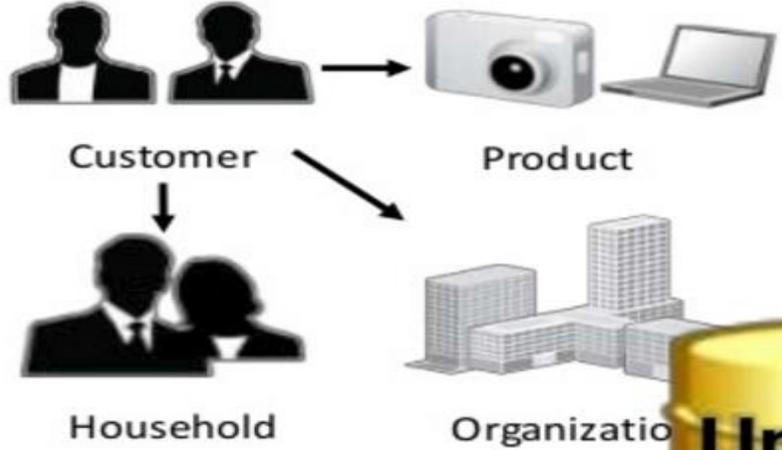
## Multi-Domain Master Data Management products

- ◆ Market direction
- ◆ Efficiencies of scale, skills
- ◆ Flexibility to use data across subject areas

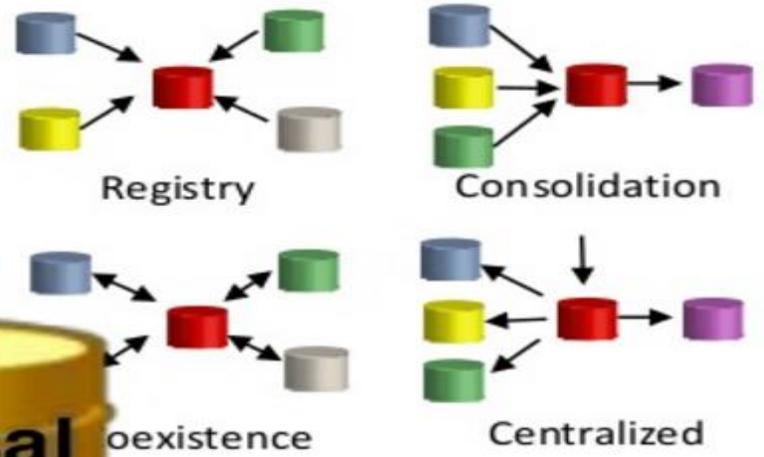
# MDM at WaTech

- Partnered with OFM on P20W Project, starting 2012
  - Design/Build Longitudinal Database (preschool through work)
  - RFP for software, including robust Identity Matching
- Enterprise Informatica toolset
  - Multi-Domain Master Data Management (MDM)
  - PowerCenter, Advanced Edition (ETL, data integration, data governance)
    - Metadata Manager, Business Glossary
  - Data Quality
  - Data Integration Hub (Real-time data integration)

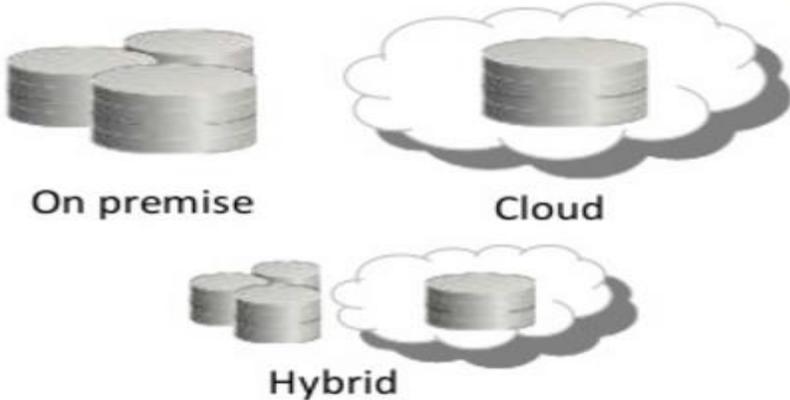
# 1 Data Domains



# 2 Architectural Styles



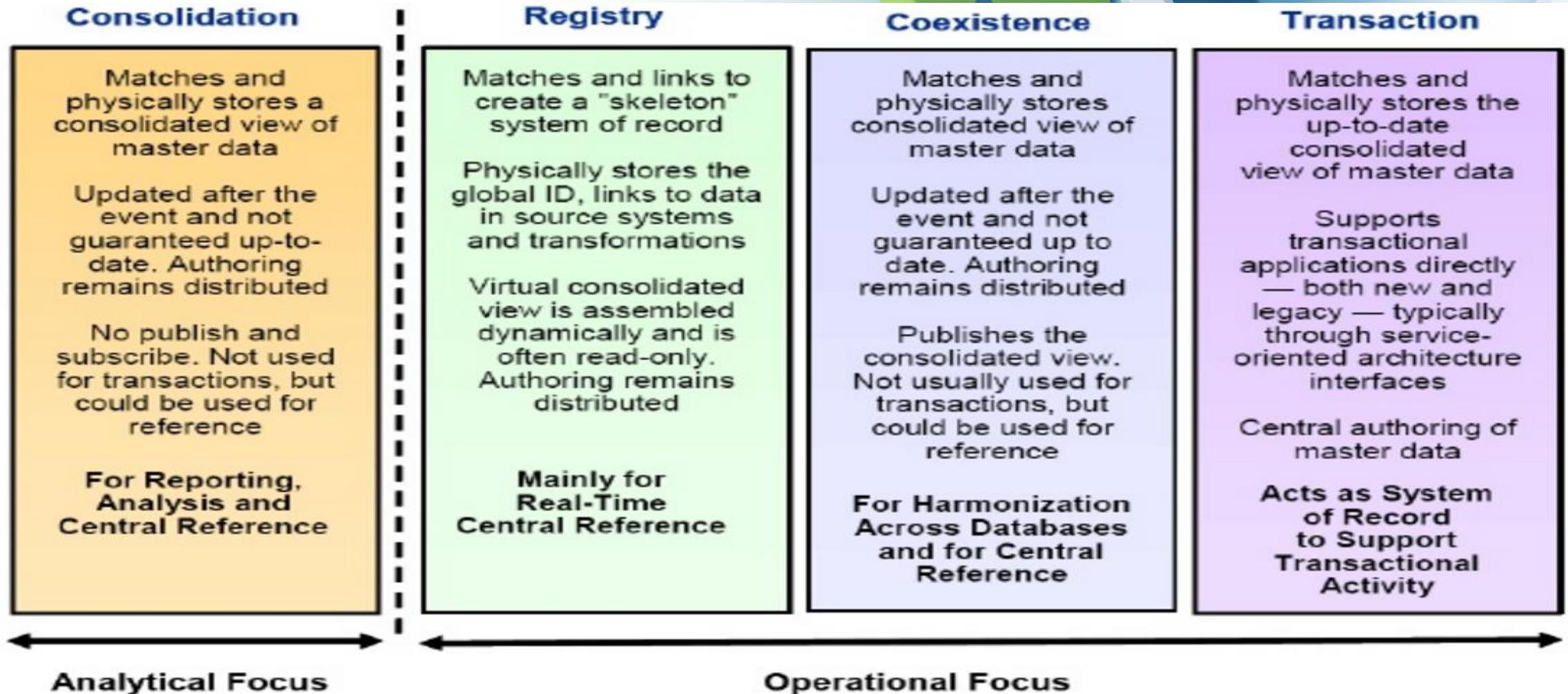
# 3 Deployment Models



# 4 Uses



# Architectural Styles



# Deployment Styles

- **Golden Record**
  - Latest, best version of each attribute.
- **Federated, Contextual Record**
  - Conflicting values maintained
- **Third Party Data Matching Service**
  - Send identifiers to service (e.g. WaTech)
  - Get back new identifiers to contact trading partner



# Should WaTech add MDM to Service Catalog?

## Potential assets to leverage:

- Software investment
- Staff training
- Production experience

## Services and approaches to consider:

- Golden Record within agency
- Client Hub for 2 or more agencies
- 3<sup>rd</sup> Party data matching service between agencies



# Big Data



**90%**  
of today's data  
has been created  
in just the last  
2 years



(...enough to fill  
**10 million**  
Blu-ray  
discs)



Every day  
we create  
**2.5**  
quintillion  
bytes of data



## Volume

Scale of data

Every  
**60**  
seconds  
there are:

**72 hours**  
of footage  
uploaded to  
YouTube



## Velocity

Speed of data

**50,000**  
GB/second



is the estimated  
rate of **global**  
**Internet**  
**traffic**  
by 2018



**216,000**  
Instagram posts



**204,000,000**  
emails sent

**1 in 3**

**business leaders**  
don't trust the  
information they use  
to make decisions



## Veracity

Certainty of data



## Variety

Diversity of data



**80%** of data  
growth is video,  
images and  
documents

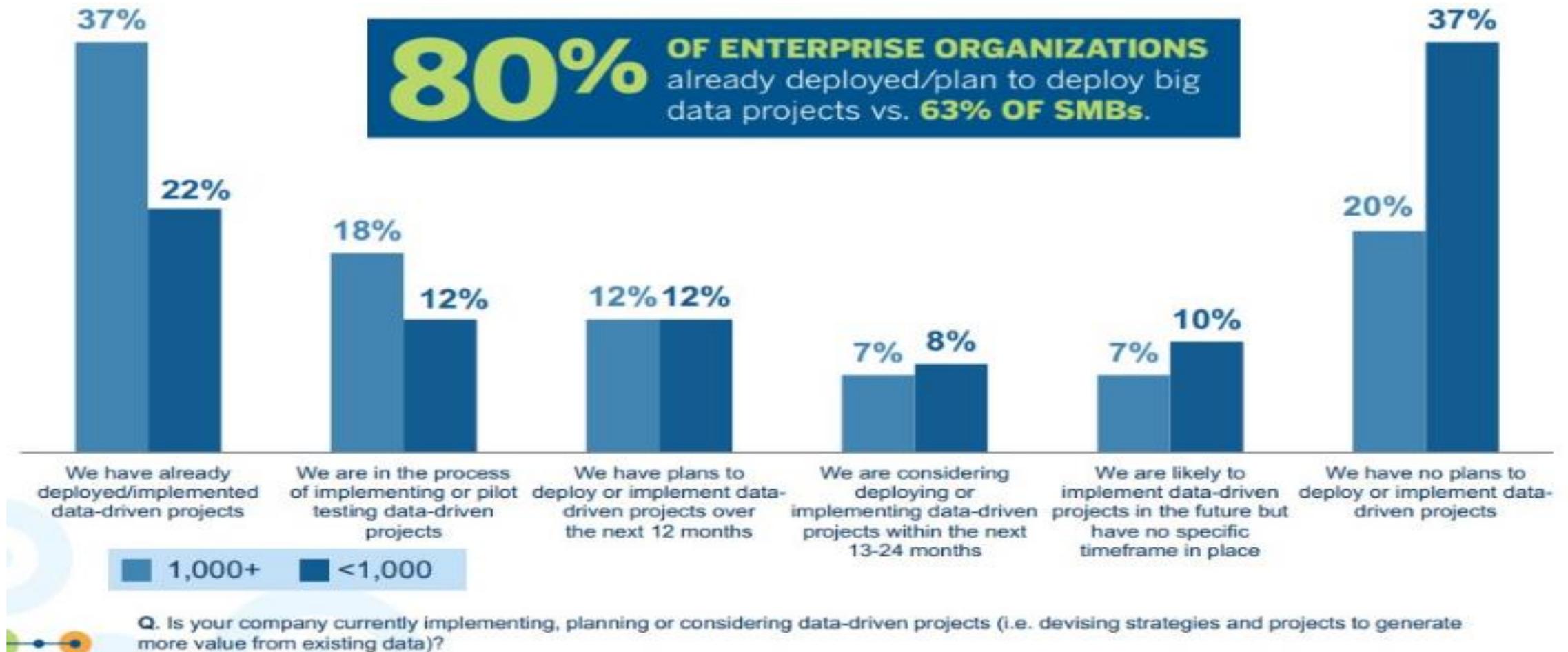


**\$3.1** trillion  
is the estimated  
amount of money that  
**poor data quality** costs  
the US economy per year



**90%**  
of generated data  
is "unstructured"  
This includes tweets, photos,  
customer purchase histories  
and customer service calls

# Industry Adoption



# Factors Driving Interest in Big Data Analysis

What data sources or challenges are driving, or would drive, your organization's interest in doing big data analysis?

■ 2015 ■ 2014

Finding correlations across multiple, disparate data sources (clickstreams, geospatial, transactions, etc.)



Predicting customer behavior



Predicting product or service sales



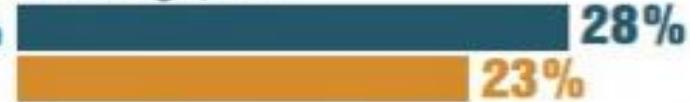
Predicting fraud or financial risk



Analyzing social network comments for consumer sentiment



Analyzing high-scale machine data from sensors, web logs, etc.



Identifying computer security risks



Analyzing web clickstreams



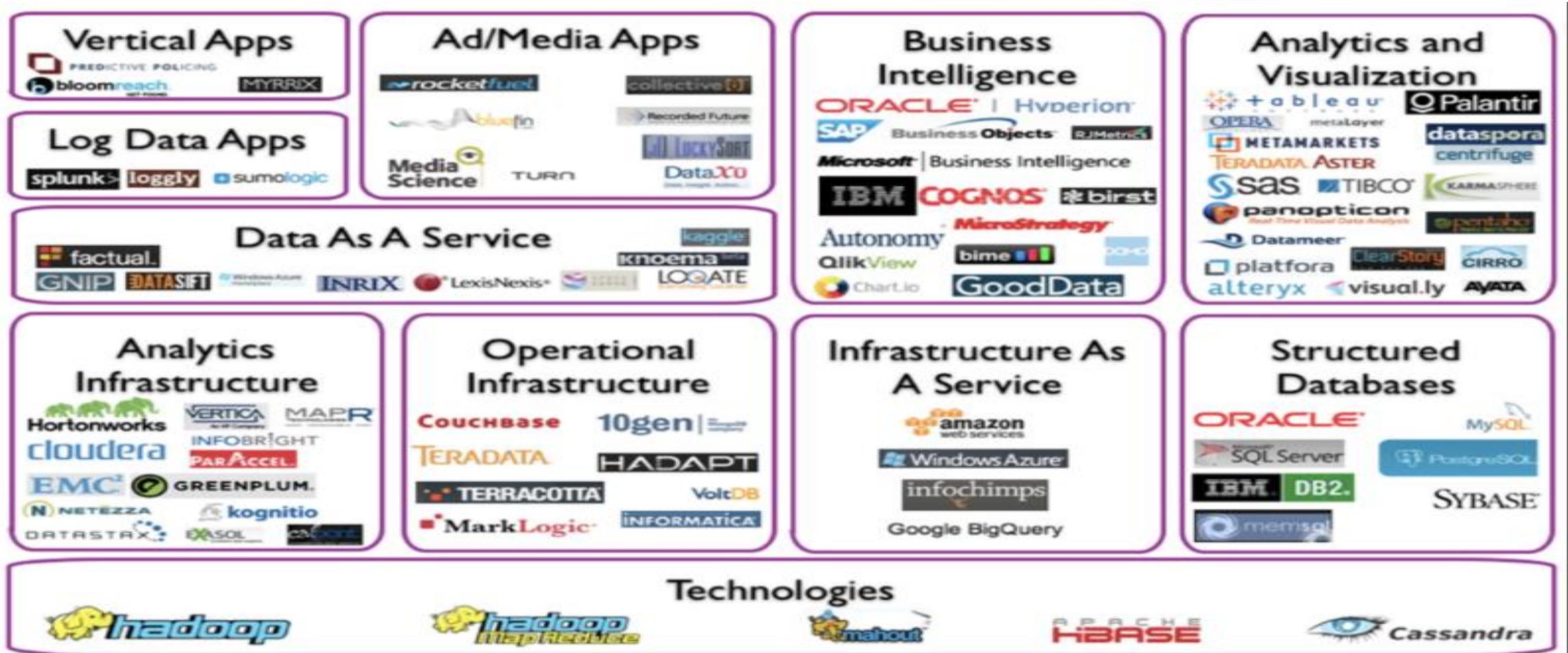
Other



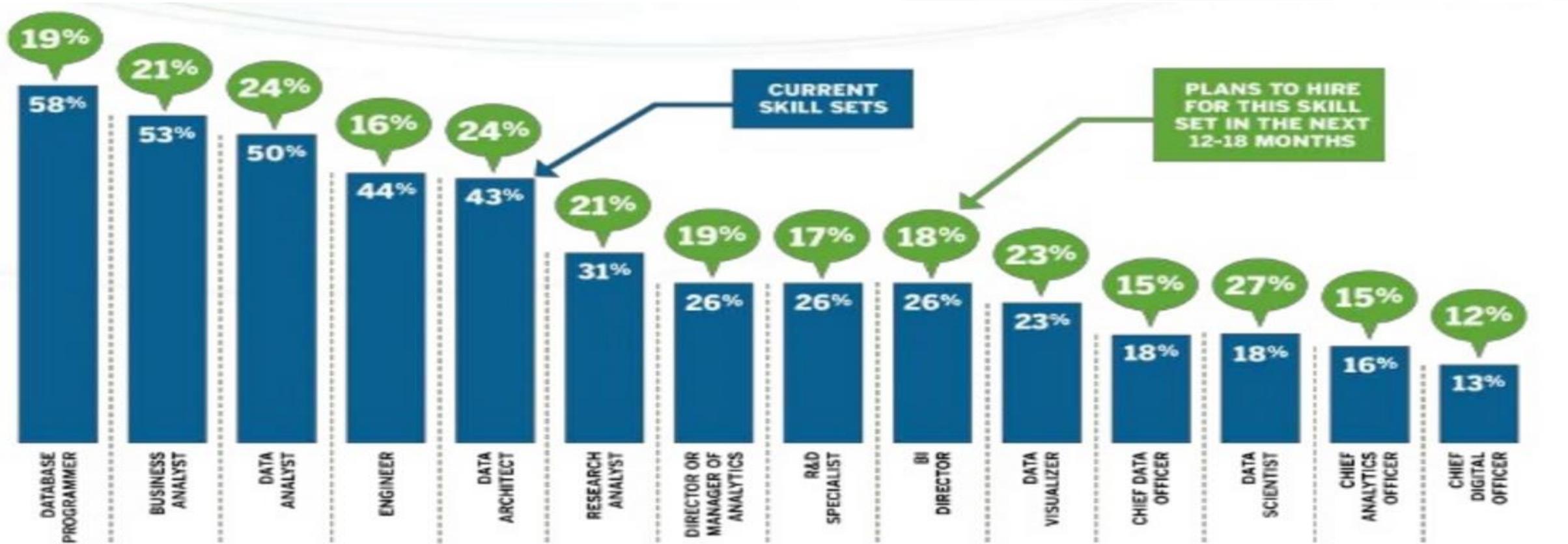
Big data analytics is not of interest to my organization



# Big Data Landscape



# Industry Staffing for Big Data Initiatives



Q. With regard to big data initiatives, what skill sets does your organization currently have? AND Q. Which skill sets is your organization planning to hire within the next 12-18 months? BASE: Plans to deploy/implement big data projects.

Source: IDC Enterprise Big Data Study, 2014

RESULTS

# Big Data & Analytics Approaches

## Do It Yourself (DIY)

- ◆ Buy or build tools, design, train, implement
- ◆ On-prem, in-cloud or hybrid

## Partner with WaTech or Other Organization

- ◆ Leverage tools, expertise
- ◆ On-prem, in-cloud or hybrid

## Outsource

- ◆ Contract to meet needs
- ◆ On-prem, in-cloud or hybrid

# Data Management & Analytics at WaTech

- **Data Management**
  - 3,000+ Online Databases (OLTP)
  - Enterprise Relational Data Warehouses (e.g. AFRS)
  - Enterprise Dimensional Data Warehouses (e.g. HRMS)
- **Business Intelligence, Reporting and Analytics**
  - Self-Service Business Intelligence (SAP, Microsoft)
  - Platform as a Service: (SAP Business Objects)
  - SAS Visual Analytics

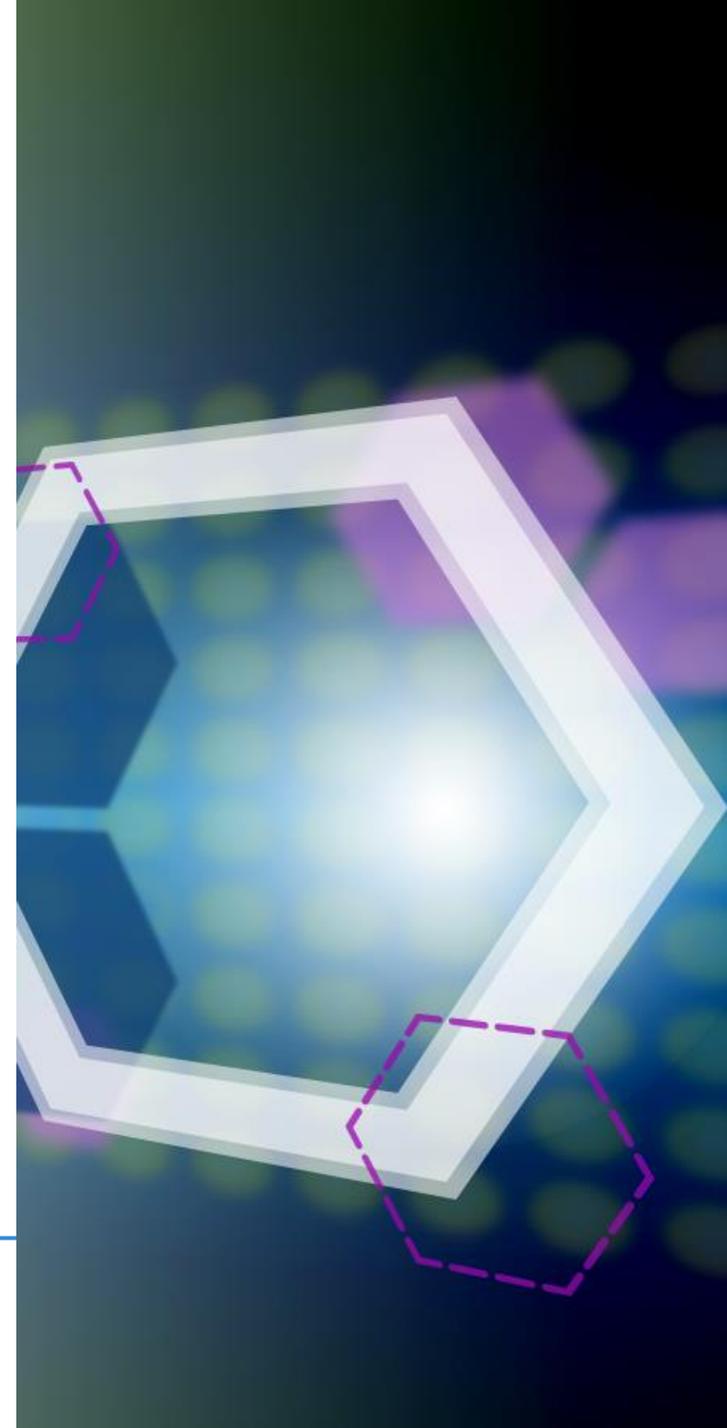
# Should WaTech add Big Data & Analytics to Service Catalog?

## Potential to leverage considerable:

- **Software investment**
- **Staff training**
- **Production experience**

## Services and approaches to consider:

- **Analytics Platform (e.g. SAS Visual Analytics)**
- **Microsoft PowerBI (extension to Office 365)**



# MDM and Big Data Initiatives in Washington State Government

## Master Data Management

- **Potential mastering of Chart of Accounts**
- **Project at Department of Early Learning**
- **AIM Project at Health Care Authority**

## Big Data & Analytics

- **PRISM tool at DSHS Research and Data Analytics**
- **AIM Project at Health Care Authority**

# Questions & Feedback



Contact: **Doug Buster**  
Data & Business Intelligence Manager

